R E P O R T R E S U M E S

ED 010 909 EA 000 156

COMPARISONS OF STUDENTS IN DIFFERENT SCHOOLS OF A LARGE INSTITUTION.

BY- HILLS, JOHN R. GLADNEY, MARILYN B.
UNIVERSITY SYSTEM OF GEORGIA, ATLANTA

REPORT NUMBER RB-5-66 PUB DATE 3 AUG 66
EDRS PRICE MF-$0.09 HC-$0.48 12P.

DESCRIPTORS- *COLLEGE STUDENTS, STATISTICAL ANALYSIS, *ACADEMIC ABILITY, *COLLEGES, *GRADES (SCHOLASTIC), ACADEMIC APTITUDE, ATLANTA

THIS STUDY TESTED THE HYPOTHESES THAT ONE SCHOOL IN A COLLEGE MAY BE EASIER THAN OTHERS IN THE SAME INSTITUTION, AND THAT STUDENTS IN ACADEMIC DIFFICULTY IN OTHER SCHOOLS WILL TRANSFER TO THE EASIER SCHOOL IN ORDER TO GRADUATE. THE STUDY WAS CONCERNED PRIMARILY WITH A LARGE SOUTHEASTERN COLLEGE WITH 12 SCHOOLS. DATA INCLUDED COLLEGE BOARD SCHOLASTIC APTITUDE TEST SCORES, HIGH SCHOOL AVERAGE GRADES, AND FRESHMAN AVERAGE GRADES OF THE 1,025 FRESHMEN WHO COMPLETED 40 HOURS OF WORK IN JUNE 1965. COMPUTER PROCESSING WAS PERFORMED FOR BOTH ANALYSIS OF COVARIANCE AND MULTIPLE DISCRIMINANT FUNCTION ANALYSIS. COMPARATIVE DATA WERE ANALYZED FROM THREE SMALLER INSTITUTIONS--A NEGRO COEDUCATIONAL COLLEGE, A 4-YEAR WOMEN'S LIBERAL ARTS COLLEGE, AND A 2-YEAR COEDUCATIONAL JUNIOR COLLEGE. A RELATIVELY SMALL STATISTICALLY SIGNIFICANT DIFFERENCE WAS FOUND TO EXIST IN GRADING PRACTICES AMONG THE SEVERAL SCHOOLS OF THE LARGE COLLEGE, NULLIFYING THE HYPOTHESIS THAT STUDENTS IN ACADEMIC DIFFICULTY WILL TRANSFER TO THE EASIER SCHOOL. SIGNIFICANT DIFFERENCES IN STUDENT ACADEMIC ABILITY WERE FOUND AMONG THE FOUR SEPARATE COLLEGES. ON THE BASIS OF GRADES IN RELATION TO APTITUDE, THEREFORE, STUDENTS CANNOT BE CLASSIFIED AMONG SCHOOLS WITHIN THE SINGLE LARGE INSTITUTIONS, BUT THEY CAN BE CLASSIFIED AMONG THE FOUR SEPARATE INSTITUTIONS. (JK)

Research Bulletin 5-66

COMPARISONS OF STUDENTS IN DIFFERENT SCHOOLS OF A LARGE INSTITUTION.

John R. Hills

and

Marilyn B. Gladney

Office of Testing and Guidance
Board of Regents
University System of Georgia
728 Hartford Building
100 Edgewood Avenue, N.E.
Atlanta, Georgia 30303

August 3, 1966

NOTE

With this Research Bulletin ends the collaboration over several years of John R. Hills, Joseph A. Klock, and Marilyn Bush Gladney. At the beginning of 1966 Mr. Klock left the University System of Georgia to enter study for the doctorate in educational research and testing at Florida State University in Tallahassee, Florida. On August 1, 1966, Mrs. Gladney left the University System to join the Personnel Department of Southern Bell Telephone Company in Atlanta to engage in selection research in the industrial setting. On September 1, 1966, Dr. Hills leaves the University System to commence a teaching career as Professor of Education in the Department of Educational Research and Testing of Florida State, coincidentally the department in which Mr. Klock is studying and from which Mrs. Gladney received her graduate degree. So, it is a small world, and we will be seeing you all again in other contexts.

Research Bulletin 5-66

Comparisons of Students in Different Schools of a Large Institution

by

John R. Hills

and

Marilyn B. Gladney [1]
Southern Bell Telephone Company

On a college campus one often finds an informal "pecking order" among degree programs, majors, or schools. Some are regarded as difficult, demanding, or prestigious while others are regarded as easy or "mickey-mouse." Underlying these attitudes appears to be the belief that a student with a specified degree of academic talent would obtain substantially higher grades in some schools or majors than in others because it is generally easier to obtain high grades in certain schools. The competition is less keen, or grading standards are lower, or some such factor produces higher grade averages for given aptitude.

Such informal hypotheses can be translated into statistical hypotheses in several ways. One translation is that with aptitude controlled higher grades by a statistically and practically significant amount will be obtained in some curricula than in others. An appropriate model for this translation is that of analysis of covariance. The hypothesis in this form states that among different schools there are no significant differences in standard errors of estimate or in regression weights for predictor variables, but there are significant differences in regression constants, and these differences are large enough to be of practical importance (Gulliksen & Wilks, 1950).

A different translation takes the form of the statement that given information on academic potential and grade performance of students, one can sort the students clearly (statistically and practically) into majors or schools. The sorting is done on the basis of the level of grades obtained relative to the level of academic potential. The model for this translation is that of the multiple discriminant function (Cooley & Lohnes, 1962). The hypothesis states that when the variables are weighted so that their combination provides the greatest separation in multidimensional space among the various groups, the degree of separation will be both statistically and practically significant. A usefully high proportion of students can then be classified correctly into their schools or majors on the basis of the application of the proper weights to the variables representing academic potential and grades.

---

[1] This study was conducted while the authors were employed by the Board of Regents of the University System of Georgia.

The purpose of this study is to determine what is revealed by these two techniques when applied to data from a situation in which the informal hypothesis is held that one school is particularly easy, perhaps even to the extent of being a school to which students who are in academic difficulty in other schools might transfer so that they may eventually graduate. The question is, are there statistically and practically significant differences among schools as revealed by these techniques, and what kinds of light do the different statistical approaches shed on the problem?

The Data

The authors obtained data from several institutions for this research. The primary institution is a large college in the southeast with one particular school out of 12 which seems to have a reputation for being easier than the others for the kind of students which the institution admits. Related to this reputation are the following facts: For an entering class which was studied in some detail, the group of freshmen who chose the, shall we say "weak", school were relatively small in number, fifth largest among the 12. They had the lowest SAT V score mean of the 12 (50 points below the mean for the total group of entering freshmen); they had the second to the lowest SAT M score mean of the 12, 64 points below the mean on SAT M of all entering freshmen; they had the second lowest high school average mean, 1/3 of a letter grade below the total group's mean; and they tied for the lowest predicted average grade of the 12 groups, 1/3 of a letter grade below the mean for all entrants. However, by graduation from the institution, this school was 3 times as large as the second largest of the 12 schools. Though less than half the freshmen graduated, 3 times as many people graduated from this school as had entered it as freshmen. It still had the second to the lowest SAT V score, third lowest SAT M score, lowest high school average, and lowest predicted average means. However, the mean senior average grade of this group was up among the rest, only 3/100 of a letter grade below the mean senior average grade for the total group of survivors from the entering freshman class. These facts seem to be congruent with the popular image of this school as being somewhat easier than other schools in this institution.

The data from this institution which were scrutinized in this study were the College Board's Scholastic Aptitude Test scores (SAT V and SAT M), the high-school average grades (HSA), and the freshman average grades (FAG), of the 1025 students who completed 40 hours of work in June, 1965, after entering the institution as first-time freshmen the preceding fall. Each student chose one of 12 schools to enter, and this choice is the classification variable used in this study.

For certain comparisons which will be discussed later, data were gathered from three additional colleges. These data were also for the persons who completed 40 hours of work by June, 1965, after entering as first-time freshmen the preceding fall. For these institutions we gathered the same data with the exception of the school of entry, which was not relevant. One of the institutions was a four-year coeducational college predominantly attended by Negroes, another was a four-year liberal arts college predominantly attended by women, and the third was a two-year coeducational junior college predominantly attended by Caucasians. The first provided 165 cases, the second provided 292 cases, and the third provided 271 cases.

The Analyses

Covariance

An analysis of covariance patterned after the Gulliksen-Wilks (1950) model was the first approach to the data. (The mathematical extensions of this model were accomplished by Dr. James Walker of the Georgia Institute of Technology Mathematics Department. The Burroughs 5500 computer program was written by Joseph A. Klock of Florida State University. It was used on Georgia Tech's computer.) Table 1 contains the N's, means, and S.D.'s, for the twelve groups of entering freshmen for SAT scores, HSA's, and FAG's. Group 12 is the school with the reputation of being weak in this institution. For this entering freshman class it has the lowest average SAT scores and has average HSA and FAG below the average for the total group of 1025 freshmen.

The Gulliksen-Wilks model provides for three tests. The first is a test of the assumption that the standard errors of estimate in the various schools differ. The Chi square for this test was 27.51, significant at the .01 level with 11 degrees of freedom. Experience with this test has led investigators to suspect that it is overly sensitive, and that this assumption is not very critical. In these data the standard errors of estimate for the 12 schools ranged in size from .38 to .57 with a mean of .49. If one is willing to ignore this degree of heterogeneity, he finds that the schools do not differ in their regression planes (Chi square = 41.36 with 33 degrees of freedom). However,

Table 1

N's, Means, and S.D.'s for Entrants of 12 Schools[a]

		SAT-V		SAT-M		HSA		FAG	
School	N	M	S.D.	M	S.D.	M	S.D.	M	S.D.
1	74	558	83	655	67	34	4	2.3	.7
2	14	529	72	613	57	28	6	1.9	.5
3	65	591	76	668	62	32	5	2.4	.7
4	104	549	78	650	64	32	4	2.1	.6
5	142	558	84	644	65	32	4	2.2	.6
6	58	532	83	616	68	31	4	2.0	.6
7	125	560	77	646	64	32	4	2.3	.6
8	28	572	93	643	71	34	5	2.5	.7
9	66	518	78	614	65	31	5	2.0	.4
10	184	559	82	650	70	32	5	2.2	.6
11	59	519	78	630	62	32	4	2.1	.5
12	106	512	79	582	65	31	5	2.1	.5
Total	1025	548	83	637	69	32	4	2.2	.6

[a] SAT scores range from 200 to 800. HSA ranges from 0(F) to 40(A). FAG ranges from 0(F) to 4.0(A).

the intercepts of the regression planes for the different schools do differ significantly (Chi square = 24.79 with 11 degrees of freedom, significant at the .01 level). Table 2 displays the intercepts in order of size from smallest

negative to largest negative, identifying the school with which each is associated. A large negative intercept implies that freshmen in this school have received low college grades relative to the level of their performance on the SAT and in high school. It is instructive to note that school 12, the school thought of as having an academically easy curriculum had the least-negative intercept. It appeared to receive grades on the average more than one fourth of a letter grade higher than the most severely graded school, group 4. While this tends to support the notion that school 12 is perhaps easier than some of the others, the difference does not appear to be very large. One fourth of a letter grade difference from the most to the least difficult schools is probably not a difference of great practical significance, and the schools with intercepts most similar to school 12, schools 3 and 8, have SAT and HSA means that cannot easily be scorned within this institution. (See Table 1.)

Multiple Discriminant Function

The multiple discriminant function technique as described by Cooley and Lohnes (1962) was applied to these same data. (The programs presented by Cooley and Lohnes are written for an IBM 709 computer. They were modified at the computer center of Georgia State College for use with that institution's IBM 7040.) The multiple discriminant function analysis yielded an F ratio of 4.7478 with 44 and 3866 degrees of freedom, significant beyond the .0005 level. Clearly then, these schools can be discriminated from each other on the basis of their SAT scores, HSA s, and FAG's with a degree of precision that is statistically significant at a high level of confidence. With only four variables, only four

Table 2

Regression Intercepts for 12 Schools, in Order of Size

School	Intercept
12	-1.40
8	-1.41
3	-1.49
7	-1.53
1	-1.54
11	-1.58
9	-1.58
6	-1.59
5	-1.60
10	-1.60
2	-1.61
4	-1.69

functions can be obtained. However, in this case three of the four functions accounted for 98% of the discriminating power of these four variables. To see whether classification on the basis of the four functions derived from this analysis was accurate enough to be of practical importance, we determined what proportion of students would be correctly classified into the schools they had entered. This was done on a sampling basis. The students within each school were placed in alphabetic order. The classifications of each of the first 10

in each school were compared with their schools of entry. Only 12 of the 120 (10%) were correctly sorted by the four discriminant functions. The predominant type of error in classification was that the small schools were assigned even fewer of the students than they deserved. Three of the larger schools, schools 4, 10, and 12, were assigned all but 5 of the students. The other 9 schools shared the remaining 5 cases. For some reason even some of the large schools did not have assigned to them their fair share of students. School 5 was assigned only one student, and school 7 was assigned none.

To avoid the influence of group size on the analysis, random samples were selected from each of the larger schools so that no school would be represented by more cases than the smallest school which had only 14 students. The 12 sets of 14 students were subjected to a complete discriminant function and classification analysis. Again the F ratio was significant beyond the .0005 level, its value being 2.0122 with 44 and 587 degrees of freedom. However, the discriminant functions in this case are not very similar to the ones obtained using all of the subjects. The first three functions for this analysis account for only 91% of the discrimination power of the four variables. This time many more students were assigned to the schools which had been undersupplied by the analysis of all 1025 students, but the proportion of correct assignments was not particularly high, only 28%.

Data With Good Discrimination

At this point one might wonder whether the discriminant function can be very effective with only four variables. We decided to apply it to data from different institutions to see whether it could classify students correctly into institutions when the institutions differed to a far greater degree in SAT scores than did the various schools within the large institution of our major concern. For this purpose we picked three additional institutions. The N's, means, and standard deviations of the four variables, SAT V, SAT M, HSA, and FAG appear in Table 3. It can be seen in that table that there are vast differences among

Table 3

N's, Means, and Standard Deviations for Entrants of Four Institutions

		SAT-V		SAT-M		HSA		FAG	
Institution	N	M	S.D.	M	S.D.	M	S.D.	M	S.D.
1	1025	548	83	637	69	32	4	2.2	.6
2	165	271	54	305	52	25	6	2.2	.5
3	292	438	89	436	82	32	5	2.3	.6
4	271	419	88	460	85	32	5	2.8	.5

these institutions in mean SAT scores, but relatively much smaller differences in FAG means. For instance, institution 1 has SAT V and SAT M means each of which is on the order of 300 points higher than those of institution 2, but the mean FAG's in the two institutions are the same. The cases of institutions 3

and 4 are of interest because SAT and HSA means are rather similar, but FAG's in institution 4 average one half of a letter grade higher than in institution 3.

The multiple discriminant function applied to these data yields the fantastically large F ratio of 362.9 with 12 and 4620 degrees of freedom. An F ratio of 3.15 with these degrees of freedom would be significant at the .0005 level.

With four groups only three discriminant functions are obtained. The first function accounted for 94% of the discriminating power of the four variables. The second function brought the percentage to 97. The first function gave great positive emphasis to high school grades and lesser positive emphasis to SAT scores, but negative emphasis to FAG. A person would tend to be at one end of the distribution of scores on this function if he had high V, M, and HSA but low FAG. He would be at the other end of the distribution if he had low predictor scores but received high college grades.

In this situation the multiple discriminant function provided accurate classification. Of the total number of 1753 entering freshman students at these four institutions, 87% would be correctly assigned to the institution they actually entered through use of these three discriminant functions. The proportions of correct assignments at the four institutions and the proportions of misclassifications from each institution to the other appear in Table 4. It can be seen

Table 4

Percents of Students Misclassified into Four Institutions

			(Institution to Which Incorrect Classification is Made)				
Correct Institution	N	1	2	3	4	Total % Misclassified
1	1025	-	0%	2%	1%	2%
2	165	0%	-	5%	7%	13%
3	292	12%	3%	-	15%	31%
4	271	9%	7%	21%	-	37%
Total	1753					13%

that very few (only 2%) of the students at the institution of our major concern would have been assigned to any of the other three institutions. Also, very few misclassifications from other institutions would be of the nature of incorrect assignments of students to institution number 2.

These data suggest, then, that the multiple discriminant function with only four variables, SAT V, SAT M, HSA, and FAG, can indeed do a highly effective classification job in a situation where there are wide differences in mean FAG's relative to the other variables. In these data wide differences seem to occur between institutions rather than between schools within institution 1. Apparently the differences between schools in institution 1 are far less extreme than are the differences between these four institutions, a fact which places the differences found in institution 1 into clearer perspective. It can be determined that there are differences among schools in grades received for a given level of aptitude in institution 1. These differences favor school 12 within institution 1 in accord with the informal pecking order in that institution. However, these differences are relatively small in the total scheme of things.

An attempt was made to determine the sizes of the differences among these institutions in grades given for a particular level of aptitude through the Gulliksen-Wilks analysis of covariance. However, in this case the standard errors of estimate were found to differ significantly, and the regression planes also differed significantly in slope. While one might choose to ignore heterogeneity of standard errors, he cannot so easily overlook heterogeneity of regression weights. Such heterogeneity implies that the planes are not parallel, that somewhere in space the planes cross each other, and elsewhere they diverge at steadily increasing rates. This being the case, the distance between the planes varies from one part of multidimensional space to another, and a difference in intercepts found at one position in space may disappear at another position. Thus no general statement can be made about the size of the differences in grades relative to level of aptitude among these institutions.

Summary

Using two techniques, analysis of covariance and multiple discriminant function analysis, this study examines the extent to which grading practices differ among schools within an institution in which one school has a reputation of being particularly easy. To provide background for comparison, differences in grading practices relative to academic potential among four different institutions are also examined. It was found that while statistically significant differences in grading practices do exist among the schools of the institution under study, the differences are relatively small compared to the differences between institutions. The classification of students among the institutions can be done quite effectively on the basis of the grades they obtained in relation to their aptitude, but the classification of students into schools within the large institution of our study was very inaccurate.

References

Cooley, W. W. & Lohnes, P. R. Multivariate procedures for the behavoral sciences. New York: Wiley, 1962.

Gulliksen, H. O. & Wilks, S. S. Regression tests for several samples. Psychometrika, 1950, 15, 91-114.